

Anopheles gambiae odorant receptor 1 genomic sequence (SEQ ID NO: 9)

5

Features:

- 1) Presumed Untranslated 5' and 3' regions are underlined.
- 2) Potential TATA box transcription initiation signal is double underlined.
- 10 3) Putative Start (ATG) and Stop (TAA) codons are in BOLD.
 - 4) Introns are tentatively assigned and are shown in lower case. Exons are highlighted.

AGCTTTGTTCATTTATGTTGAAATCTAGCCCATTTTGTATAGTGCTGAACGACGAAGAACATACGAAAGTACCTCGT 15 CCGAACACTATCAACATTAATTATACCAAGCTAGAAGAAGATATTTATAGTCAAGCCTCAACATCATAGGAAACTTT AGCAAAACCATTTAATTTACATGATGATAAGTCCCACCTCTTACCCCAGCACAGGTTTGAGAAGGACGAAAGTATCT $\tt TTACGATAATATTACTCTAAGGTAGTTTTTGAATAAAATAAAATTTACGTGCAAGTGGTGGCATCGGACATCA\underline{TTC}$ $\underline{GAACACGTCAGGACATAACTGCGACATGCGTATGGTCAGT\underline{TCCAC}TAGTGCCAACACTGGTTCCAGGGCACTACCTT}$ 20 ${\tt CCGAAGCAGTAGAACCTAATGTATTGGAAATTATTAGGACATACTGCAACATGCATATGGCTAGTTCCGCTGGTACC}$ AACGATGGCACCAGGACACTATCTGCGGCCTTGTAAAATCACTGTAAAATCTATACAAAAACGGCTTTACCCATACT TTATCACAAAAACGGCAGGTGAGGGCTGGATTGCTTCAAAGCATTAGAAATATATAATTTCAAAGTCCATAATCTCC AATGCTCAATTGTTGTAGATTCGTTGGATGACTCTCGCTACGTGCTATAGTGGTCAATACTTCCAATTAGATTTCAT 25 AATTAGTTTCCAATTGTCCACGGAAAACCCaCAAAAGAAAAAAAAACTTGTATCTAGGGTGGAATTTTTCGAGAACA attctccaaattctgcagaataattctgcaaattttacaaaactgctcaaccaccaataattccaattaatcatctg aacatttaaaactgataattaagatgagtaattgcttcgtcatcacctaagaaatcgattagtttggataaaaagaa 30 aattcctcgttgaaaattggtctcctatagttctgctaacgggccacttcaaaagcaagaactaacaaaatcataat tatggtgcaagtaactatcagtaccagtaatcgccattaaaaacttttcctcaatttgcggctcgttaccggctaaa tacagagcagagtaacgggaagtgatcaacgtcgctattagtataacgaggaacgccctccgaaggtgtgttgaagg accttttcaaattgaaaccaagtactgtttccagttttaaattggatagttataaaatgagccgttcaacgatcggg catcatttgagtttcatcttcgaggagaaatagatcagtgccactgtttaaccgaaagtaatgaagctgaacaaact 35 GGCCACCGGAAGATACGGATCAGGCAACGCGGAACCGGTACATCGCGTACGGTTGGGCTTTGCGGATCATGTTTCTA CATCTGTACGCTCTAACGCAAGCCCTATACTTCAAGGATGTGAAGGATATTAATgtgagtctctagttagctattag tgttccacctgtccataatctgtcttttattgggtagGACATCGCAAATGCATTGTTCGTGCTTATGACTCAAGTGA CGTTGATCTACAAGCTGGAAAAGTTTAACTACAACATCGCACGGATTCAGGCTTGTCTGCGCAAGCTTAACTGCACA 40 CTGTATCACCCGAAACAGCGCGAAGAATTCACgtaagcctgctgggaaatatgactaaaaagagtgctaacaaacga ctctcctccaaatgtag<mark>ccccGTTTTACAATCGATGAGTGGAGTGTTTTGGCTGATGATCTTTGTCATGTTTGTGGC</mark> <u>FATCTTCACCATCATCATGTGGGTTATGTCGCCAGCCTTCGACAATGAACGTCGTCTGCCCGTGCCGGCCTGGTTCC</u> CGGTGGACTATCACCATTCGGACATAGTGTACGGTGTACTGTTCCTGTATCAAACCATTGGAATCGTCATGAGCGCA ACGTACAACTTCTCGACCGATACCATGTTTTCCGGCTTGATGCTACACATAAATGGACAAATTGTGCGGCTTGGTAG 45 ttatcagCTTGGACATGACGTCCCTCCCGAACGCCAATTGGTCGCAACGGATGCGGAATGGAAAGAGATGCGAAAGC TTAAGqtacqaattgggccaattaattgtgtcatttaaaaagcttgacccaacttttcacagcttcggcgatgaagt gcaggacattttccaagGATCTATCTTCGCGCAAGTATGCGCGTCTGTAATTATCATTTGTATGACACTGCTGCAAG 50 TACCGGGGGCGATGTTACGATGGCCGATCTGCTGGGCTGTGGGGTCTATTTGCTAGTAAAGACATCGCAAGTGTTTA TTTTCTGTTACGTAGGGAATGAAATCTCCTATACGgtaggttggacacgtagaggaattaaatgtttgggaagaata tcaataccaaatagtatgatgtttcgttacagACGGATAAATTTACAGAGTTTGTTGGGTTTTCCAACTACTTCAAG TTCGATAAGCGTACCAGCCAAGCAATGATATTTTTTCTGCAAATgtgagatagcggtgtatttgtgcagtcagtaca

ttaaatacgttctctatttcageachenaacheneacheneacheneachaechaechaechaechaechaechenea

Figure 1 continued

15

20

Anopheles gambiae odorant receptor 2 genomic sequence (SEQ ID NO: 10)

5

Features:

- 1) Presumed Untranslated 5' and 3' regions are underlined.
- 2) Potential TATA box transcription initiation signal is <u>double underlined</u>.
- 10 3) Putative Start (ATG) and Stop (TAA) codons are in BOLD.
 - 4) Introns are tentatively assigned and are shown in lower case.
 - 5) Exons are highlighted

GGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTCCCTCACCGTGACGTGCTAGAAATGGTTCAACATACTCGT 15 CCGGCAGAGCGAAGACGACGAACAGCGGAATGTCCCAGGAAATGTAATGAGATATCACAGCAAGTGAACCCAAACCG AAATCCACTGACCACTGGCCACACATCAACCACCGGAGCGGGGGCCTCAGTGCCCAGCGAAGC<u>ATATAA</u>TTTGCTCA AAAAGTCACGGTACTCAATTAATTTGATTATAATCAATTTCGTGGCTTCCAACACCCCTTCTTCCACAATCCATCG CCGAGTGAGCGAGTATAAAGGTGAAGAAACGTACCTTGCGCTTGCTCACTAACTGAACCGGATTTCAAAAAGGAACA 20 TAAACCGCAACCCACAGCCGAAA<mark>ATGCTGATCGAAGAGTGTCCGATAATTGGTGTCAATGTGCGAGTGTGGGTGTT</mark> ACCTCCTCC tacgtgggcgagggggggggcaataaccttcccacttggtggatattttcataccttttccatgtgtt tttttattctctgtttgttgccatccagerecaacerecarrerecarearcaareeaeecaaarreaeaerecarearrere 25 NTGAAGGCGTTGGGGGGGGAGTAGGGTGTGGAGGtaagtcattggtttttctagtttttgggggagttgtttaca ccataaccaccccgacggtaacatttgatcgtcccgcgaaaatgtttgtacag<mark>AAAAAAGGAGACAGGGAGAGGGGI</mark> 30 <u>eAeAWAeAWATEATEGAE</u>taagtagacgctagtagactcgaccggattgcccttccctcggggaggggaggtttgct atttcgggatgcggcagcacgcatacacacacacaggaagccattaattctcccgttttcatgcccgcacgggcact gggtcatgtttcacatccttcctttccaaacacacacacgcgcgtgcacgtacagATATGTTCATCAT 35 AAGCATTGtaagtaaaatcgaccgacgtgcggtcgctagtccgtctccggactctcatttcgggactcaatcgttcc tagatcggctgtcttacattgttgtgtttctgcatggggatcggttttgtttttcctctccatttcagAGCCTAGGC 40 ANHIGE CENTRECONTRIPINACIAN RECENCE CERCE CECETA CENTRECA CENTRACE CANDANCE CANDANCE CHURCHUNAN CHUNAN CHURAN C ACCINCENCIACCEACCEANCE RECEIVACE tttggctgatcgatgctctgttcaatgaacatggcacagaaggctgtgta aatagctgttcattaataagttttttcagaatgtatcgtttttagttgatttaaacgcattgttctatgcaatggta 45 50 ttattattattattgttattattattcttattattgctattgttattattattattattcttattgttgtt

acaataatctctaagaattaaaattgcattttgtaatgaaatatgttgattgttcgaatagttcagaaaaacttaaa aatgcctcagcattaaacagttttgaggttgttcagggcatttagtttagatattttagtattttaaagcatttgtt ttcattactacaaaaaagcaaatttatgagtgaattactttcagttcttctaaacgcctatgtgtatgcaattacat 5 aacaataqctctcttttttattqcatttttccttaqtaatctaaatccaatctcttctttccctcttgcagATTAAA EHICECCAYACONGHICHIYACOCCCAAHCCAACCHHILCCAAAAAHCHHILCCAAAAAAAHHLCCHICCAAAACCHICCHLCCHLAACHCCHAAHHHILCAAAAAAAHHLCCHICCAAAACCH GATCATCAAACACCATTAGCAGCCACAAAGTTACCAGCCGCTTATCCCACGGGATTTGGTGGAAAGTTATTGCACTG 10 ${\tt GCGACGGTGAAAAAACGCTGCATTATTGTGCTTGCTTCAGCATTCCAGCGAATGACTCTTAAACTTTTCCATTCAAA}$ AGTCGCGATGCTCACGATACGGAGCGGTGTGTTGTTCGATCCGCCGAGTGCACTCGCAAGCCGGTGATGTTGCCGGT AAGGTCTCTGCTCCGGGGCATGGATTCTTTCCCCCTCCGGGTGGTTGGGGGGTATTGTTTAGGTTTTTATTTTACAAA 15 ACGAACATGGCCAACAAACACAGCTTCTATCTCATCTCTGTGTCGCACTGTCTCGCTTTCCCGCTGCGTTGCTTGTA <u>GTACTATCATTGTTTTAGTCCACGGGTTT</u>ACTTCTAATTCCA<u>TTGCAC</u>CACGCAAAAAGGCTCATCCTTTGCTCGTT 20 GCTGTGTGCGCTCGAGTCAGCCGACGGTACAAGGTTTAACCGGTACAAGCAACTCCCGGACCGATCCCAAAACTCTG AAGAGCGAGAAACATTGGTACGATTTGGTGTGGTTAGCAAATTTGATTTCCACTGATTTTGAGTGCAAATTTAATGC ATCGAAAATTTGCCATTCAGGGTAAAGTTGCTCGTGGACGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTAT 25 CGATACCGTCGACCTCGAGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGGA

Figure 2 continued

30

Anopheles gambiae odorant receptor 3 genomic sequence (SEQ ID NO: 11)

5

Features:

- 1) Presumed Untranslated 5' and 3' regions are underlined.
- 2) Putative Start (ATG) and Stop (TAA) codons are in BOLD.
- 3) Introns are tentatively assigned and are shown in lower case.
- 10 4) Exons are highlighted.

AAGCAGAACACATCAAGAAGCAATTAGGTGTGTCGTACGTTAGCAAGTAGTTCGCGAGGAGGAATAAAATAG**ATG**CC TTCTGAGGGGCTTCGTCTCATTACTTCCTTCGGAACTCCTCAAGACAAACGCACGATGGTACTGCCAAAATTAAAGG ATGAAACAGCAGTGATGCCGTTTCTGCTGCAAATTCAAACCATTGCCGGACTGTGGGGTGACCGTTCCCAGCGGTAC 15 CGTTTTTATCTCATCTTTTCCTACTTCTGCGCGATGGTGGTTCTACCCAAAGTGCTGTTCGGTTATCCAGATCTCGA GGTTGCGGTACGCGGCACGGCCGAGCTGATGTTCGAATCGAACGCATTCTTCGGCATGCTAATGTTTTCCTTTCAAC GCGACAACTACGAGCGATTGGTGCATCAGCTGCAGGATCTGGCAGCTCTAGgtgagtatgcagccaatcgattgttc caaaccttcgcaacatccttcgtaacactgctacactttcagTCCTCCAAGACCTACCCACAGAGCTGGGAGAGTAC CTGATCTCAGTGAACCGACGGGTCGATCGGTTCTCCAAAATTTACTGCTGCTGTCACTTTTCCATGGCAACGTTCTT 20 TTGGTTCATGCCCGTCTGGACGACCTATTCCGCCTACTTTGCTGTGCGCAACAGCACGGAACCGGTCGAGCACGTGT TGCACCTCGAGGAAGAGCTGTACTTCCTGAACATTCGGACTTCGATGGCGCACTATACGTTTTATGTGGCCATTATG TGGCCCACGATCTATACGCTCGGGTTTACCGGTGGCACAAGCTGCTGACCATTTTCAGCAATGTTAAGTACTGTTC GGCCATGCTGAAGCTCGTTGCACTCCGAATCCACTGTCTAGCGAGAGTAGCGCAAGACCGAGCGGAAAAGGAGCTGA ACGAGATTATTTCCATGCATCAGCGGGTACTCAAgtaagtaaattcaaattgaaagttttgcagggaataacttgag 25 tgtgtctgacccgtgcacatcctagCTGCGTGTTCCTGCTGGAGACGACATTCCGCTGGGTATTTTTCGTGCAGTTC ATTCAGTGTACAATGATCTGGTGCAGTCTCATCCTCTACATAGCGGTGACGgtaatagcattttcgtcatttcgtta gccttattcaatccatttttgtgaacgtgaatttcccccagGGGTTCAGCTCGACGGTAGCGAATGTATGTGTCCAG ATCATTTTGGTGACGGTGGAAACTTACGGCTACGGCTACTTCGGAACAGATCTAACCACGGAGGTGCTTTGGgtacc ctttggatgaagcttcaaaaagtaattccaaattctgttttcgatttttccccttttccactagAGCTATGGCGTTG 30 CCCTCGCCATTTACGATAGCGAGTGGTACAAGTTTTCCATTTCGATGCGCCGCAAACTTCGACTGCTACTGCAACGA TCCCAAAAACCGCTCGGCGTAACGGCGGGAAAGTTTCGCTTCGTCAATGTGGCCCAGTTTGGCAAGgtaacattaat tacagtttgaaaattctgaagaatgcatcttacttgccttacttgttgttccagATGCTCAAGATGTCCTATTCATT TTACGTAGTACTGAAGGAGCAGTTTTAGGAGCTGCTGTTTCCCACCCTGGAAATGGCCTTTTCGCACTGTCTTCTGT 35 ACAGCTGAAGGACAGGGTACAATTTTTGCTGCTGTTATTACGCGCAGCGCATTGGATACGAAAACATTGGCCACAAG TTCTACGATTTTAGCGTTTATTTACTGTTCGTAGCAGCTTTTTTCCaCAATAAACACACACAATAACGTACCGACAG **ACGA**

40

Anopheles gambiae odorant receptor 4 genomic sequence (SEQ ID NO: 12)

5

Features:

- 1) Putative Start (ATG) and Stop (TAA) codons are in BOLD.
- 2) Introns are tentatively assigned and are shown in lower case.
- 10 GGGGAACTCCCCCACCGACCAGACGACGAAAGCTAACGATGTGCAATTGAATAGTCATTAGT AGCGTTTTTGCTCGCAAACGAACTAACCCTTTGACTTTTTAAGTTCACTACGGTGAGGACAAAAA CGTTCCATCGACTACATAATCATAATTATATGCCACATTTTATTATAAGTTTTTTGTATCATTTTTA AACAACACAAAAATGCATCCTTTCGAATATTAGTCAGGTTGTATCAACAATGAAGTTTGAACTGT 15 TTCAAAAATATTCCTCCCCGGACACGGTCTTATCCTTCGTGCTAAGGCTTTTGCATATCGTGGGC ATGAATGGGGCAGGATTTCGGTCGCGAATTCGAGTTGGTGGCATTTTTCTGTTCTATTTAATCTT TCTTGTAATACCGCCACTAACGGGCGGGTACACCGATGGTCACCAGCGTGTACGCACCAGTGTG ${\tt GAATTCCTGTTTAATTGCAATATTTACGGCGGCAGTATGTTCTTTGCCTACGATGTGGCCACTTT}$ 20 GAGCGGATATTATCGCCAAAGTGCAAACGACCTGCATGGGTGCTGTAACGCTTTTCTACTGGAT ${\tt TGCACCGATACCTTCCATCTGTGCGCACTACTACAGGTCGACCAATTCCACCGAACCCGTGCGG}$ TTTGTGCAACATTTAGAGGTGAAGTTCTATTGGCTCGAGAATCGCACCTCAGTCGAGGACTACAT AACCTTCGTGCTGATCATGCTACCCGTCGTGGTTATGTGTGGTTACGTATGCAATTTGAAGGTGA 25 TGACCATCTGCTGCAGCATTGGACACTGTACACTGTACACCAGGATGACTATAGAGATGGTAGA GCAGTTGGAAAGCATGGCATCAGCGGAACGAACTGCCAGCGCCATACGCAACGTGGGGCAGAT 30 GGTAATGTTTTTTTTTCTTGCCACTGCGGAAACTTTCCTGTATTGTTTACTTGGGACGCGGCTTGCGA CACAACAGCAGCTGCTGGAGCACGCACTCTATGCTACACGGTGGTACAACTACCCAATAGCCTT TCGCAGCAGCATTAGGATGTTGAGACAGTCGCAAAGGCATGCACACATAACGGTGGGGAAG TTTTTTCGCGTTAATTTGGAAGAATTTAGCAGGATTGTCAACTTATCCTACTCTGCTTACGTCGT AGTTTTCCGAATCTATATTAGATCTAGAATTTAATCTAGATGTCATAATATGATCTTGGCCATGA 35 CCGGTTCCTGGTTTTGGAACCAATTCTCAAAACAATTTTGAACTTAGGGCGAGGCATGAAATGTC CCAAGAACCTATCCAAGTTCTGGAACTACATATTACCGAATCTATCCCATTATTGCCTCGGAACT GGTTTGGTGCTAAATATTTGTCCAAATGTTGGTCCTGGACCTATCCAGACAAAGATCTTCAATTA
- CCGGTTCCTGGTTTTGGAACCAATTCTCAAAACAATTTTGAACTTAGGGCGAGGCATGAAATGTC
 CCAAGAACCTATCCAAGTTCTGGAACTACATATTACCGAATCTATCCCATTATTGCCTCGGAACT
 GGTTTGGTGCTAAATATTTGTCCAAATGTTGGTCCTGGACCTATCCAGACAAAGATCTTCAATTA
 TTCCTACCACTGGAACTGATTAATTGATGTAGGAAGTCATGGAGGTGTTCAGGGAGAATTTAAA
 40 CACTAATGTTCCAACTCATTATTTCAAGGGCAATTCTATTTTTATATGCCCCTACGGATTGATAC
 GTATGTATTACTCCATTTCCTGGACTTTGTCTTATTCTTGCTGCTGATTGGACGTGAAATGTTGA
 GAAAAAGATTCTTATTTATGAGTGATACAGAGCCTTTAAATACTCCTACGTTGTTTTGCTATTTAA
 GTATGGCCAGGCTAATCACAATCGCTACTAATGAACAGAATCTCTTCTAATTAAACCCCTTTCGAT
 TGATAGTGTCAATGTCAATGTCGAGATAATTGAACTGCAAACgATACCTACCTTAAACGGAGCAG
- 45 AACACATCAAGAAGCAATTAGGTGTGTCGTACGTTAGCAAGTAGTTCGCGAGGAGGAATAAAAT

Figure 5 ANOPHELES GAMBIAE

Preferred DNA Codons

Amino Acids			Preferred Codons					
Alanine	Ala	A	GCC	GCG	GCT	GCA		
Cysteine	Cys	C	TGC	TGT				
Aspartic acid	Asp	D	GAC	GAT				:
Glutamic acid	Glu	E	GAG	GAA				
Phenylalanine	Phe	\mathbf{F}	TTC	TTT				
Glycine	Gly	G	GGC	GGT	GGA	GGG		
Histidine	His	H	CAC	CAT				
Isoleucine	Ile	I	ATC	ATT	ATA			
Lysine	Lys	K	AAG	AAA				
Leucine	Leu	L	CTG	CTC	TTG	CTT	CTA	TTA
Methionine	Met	M	ATG					
Asparagine	Asn	N	AAC	AAT				
Proline	\mathbf{Pro}	P	CCG	CCC	CCA	CCT		
Glutamine	Gln	Q	CAG	CAA				
Arginine	\mathbf{Arg}	R	CGC	CGG	\mathbf{CGT}	CGA	AGA	AGG
Serine	Ser	S	TCG	AGC	\mathbf{TCC}	AGT	TCT	TCA
Threonine	Thr	${f T}$	ACG	ACC	ACT	ACA		
Valine	Val	V	GTG	GTC	GTT	GTA		
Tryptophan	Trp	W	TGG					
Tyrosine	Tyr	Y	TAC	TAT	, 1 ,			.,

http://www.kazusa.or.jp/codon/cgibin/showcodon.cgi?species=Anopheles+gambiae+[gbinv]

Name	SEQ ID NO
Arrestin 1 (cDNA)	SEQ ID NO: 1
Arrestin 1 (polypeptide)	SEQ ID NO: 2
Odorant Receptor 1 (cDNA)	SEQ ID NO: 3
Odorant Receptor 1 (polypeptide)	SEQ ID NO: 4
Odorant Receptor 2 (cDNA)	SEQ ID NO: 5
Odorant Receptor 2 (polypeptide)	SEQ ID NO: 6
Odorant Receptor 3 (cDNA)	SEQ ID NO: 7
Odorant Receptor 3 (polypeptide)	SEQ ID NO: 8
Odorant Receptor 4 (cDNA)	SEQ ID NO: 13
Odorant Receptor 4 (polypeptide)	SEQ ID NO: 14
Odorant Receptor 5 (cDNA)	SEQ ID NO: 15
Odorant Receptor 5 (polypeptide)	SEQ ID NO: 16
Odorant Receptor 6 (cDNA)	SEQ ID NO: 17
Odorant Receptor 6 (polypeptide)	SEQ ID NO: 18
Odorant Receptor 7 (cDNA)	SEQ ID NO: 19
Odorant Receptor 7 (polypeptide)	SEQ ID NO: 20

Anopheles gambiae odorant receptor 5 genomic sequence (SEQ ID NO: 21)

Predicted Exons: ITALICIZED, UNDERLINED AND HIGHLIGHTED.

Introns: lowercase.

5

10 tctagacttgaacccatgacgggcattttattgagtcgttcgagttgacgactgtaccacggggaccacccgtttatcactattaattaattataatatgettttgtagegateageetacegggttttgtttetetggatatettaagtteeeatttgattateaagatagaa caacaacttgtaccttaaataatcattacgtacccttaatcaacctgtgcatcaaggagttttcgcgaaagcaaaaatccgattgtct gatgttgtcttgattccatccgattcgttactggttctgcaaaatcgtccaataatacggcaatgtccttatcgatgcttgaatcaacat cacattgtttgcatttcgttttttgcgtgcaaatatgttatttgcaaagaaggcaaggtaatgtgcttaagagtaaatacaattcgctg 15 TGCTACCGAAGCTGTCCGAACCGTACGCCGTGATGCCGCTTCTACTACGCCTGCAGCG TTTCGTTGGGCTGTGGGGTGAACGACGCTATCGCTACAAGTTCCGGTTGGCATTTTT AGCTTCTGTCTGCTAGTAGTTATTCCGAAGGTTGCCTTCGGCTATCCAGATTTAGAGAC*AATGGTTCGCGGAACAGCTGAGCTGATTTTCGAATGGAACGTACTGTTTGGGATGTTG* 20 CTGTTTTCTCTCAAGCTAGACGACTATGATGATCTGGTGTACCGGTACAAGGACATA AAAGATTGgtgcgtgataatgattgataaaaggaacctttgagcaactcctatccctttcaagCTTTCCGTAAGGACCAAGATCTACTGCTGCAGCCATCTGTGTTTTGGCCATCTTCTACTGGGTGGCTCCTTCCCAGCACCTACCTAGCGTACCTGGGGGCACGAAACAGATCCGTCCCGGTCGAACATGT 25 GCTACACCTGGAGGAGGAGCTGTACTGGTTTCACACCCGCGTCTCGCTGGTAGATTACTCCATATTCACCGCCATCATGCTGCCTACAATCTTTATGCTAGCGTACTTCGGTGGACTAAAGCTGCTAACCATCTTCAGCAACGTGAAGTACTGTTCGGCAATGCTCAGGCTTGTG GCGATGAGAATCCAGTTCATGGACCGGCTGGACGAGCGCGAAGCGGAAAAGGAACTGA TCGAAATCATCGTCATGCATCAGAAGGCGCTAAAgtaaggtctgccggtatgttgtggatagaatacattt ctagetgettteagATGTGTGGAGCTGTTGGAAATCATCTTTCGGTGGGTTTTTCTGGGACAG 30 TTCATACAGTGCGTAATGATCTGGTGCAGCTTGGTTCTGTACGTCGCCGTTACGgtaacta aaagcactgtagtgatctgtctgccacaccattcactgctgtgtcttgttttgtcactcttcccag/GGTCTCAGCACAAAAG CGGCAAACGTGGGTGTACTGTTTATACTGCTAACAGTGGAAACCTACGGATTCTGCTA CTTTGGCAGTGATCTTACCTCGGAGGCAAGTTGTTATTCGCTGAgtttcagttacttttccgttcccc tctaaccgtaccacttgtaccatttgtttgagacagagcttgagcgtagCACGTGCTGCGTACGGTAGCCTCTGG 35 TATCGCCGTTCGGTTTCGATTCAACGGAAGCTTCGAATGGTACTGCAGCGTGCCCAGA *AACCGGTCGGCATCTCGGCTGGGAAGTTTTGCTTCGTCGACATTGAGCAGTTTGGCAA* TetatggggagaccttccactgtggcaagaaagattttctttattaatgcatcttttaatttacagATGGCAAAAACATCA TACTCGTTCTACATCGTTCTGAAGGATCAATTTTAAagggggaactcccccacccgaccagacggaa 40 agctaacgatgtgcaattgaatagtcattagtagcgtttttgctcgcaaacgaactaaccctttgactttttaagttcactacggtgag cataatcataattatatgccacattttattataagtttttg

Anopheles gambiae odorant receptor 6 partial genomic sequence (SEQ ID NO: 22)

These are the predicted last three exons of another candidate *Anopheles* gambiae odorant receptor.

Predicted Exons: *ITALICIZED*, <u>UNDERLINED</u> AND <u>HIGHLIGHTED</u>. Introns: lowercase.

10

ttgatgccgtatgcgccgcgtgctataggctag/TTATGCTTACCGGATGTTGCGATCGCGCACGTGCTTT TCCGCATACGCCAGTGCACACTTGATGGCGGTGGTGATGACGTCTGCTGCGCACCGTT 15 ACAGACGGTTAGACGGATATATGCTGGTAAAGTTTGTCCTCTTCATGCTGTGCTTTCTG ATCGAGCTGCTGATGCTGTGTGCGTACGGTGAGGATATTGTGGAATCGgtaaggcaccaggc ggtgatgagcgagtcgcgagtaattgaagcttttgcttttaaaacacatcagagCCTTGGGGTGATTGATGCCGCT TACGGTTGCGAATGGTACCGGGAAGGGTCGGTGGCGTTCCATCGATCCGTGCTGCAAA 20 TTATACACCGCAGCCAGCAGTCCGTCATACTGACCGCATGGAAAATTTGGCCCCATCCAA $\underline{ATGAGTACTTTCAGTCAG} \\ \textbf{gtgagttgccaattgattgccgtttgcgttaatatttcagtaagagtgcgctctttcccttag}$ ATCCTGCAAGCTTCCTGGTCCTACTTTACCCTCCTGAAGACCGTCTACGGGAATAAgtaa gcgcgagagagagagagagagagtatcgttcaccctttggatgaatcaatagatttctaatcatgaaccattgaaaaatgaatca acattttcgctagttgcacaatattgtaccattctatacagcttcaccacgaccaagcgtttgttgcatcaggaccaaacacgtttcga25 caagccgcgtcacctgctggc

Anopheles gambiae odorant receptor 7 genomic sequence (SEQ ID NO: 23)

5

Features

- 1. Predicted Exons (7): ALL CAPS, ITALICIZED, UNDERLINED, HIGHLIGHTED
- 2. Introns (6): lowercase
- 10 3. 5' and 3' sequences: lowercase, dotted underlined

ccgcccgggcaggtgacttacgcggtctgacttgctggtgcgctgctttgtacggcaaacggctacacaagcgaatcgaattattttcc15 gcattgtgtttagtgagaagtgaaaagaaaagtgctgaaaaatgcaagtccagccgaccaagtacgtcggccttcgttgccgacct gatgccgaacattcgggttgatgcaggccagcggtcaactttctgttccggctacgtcaccggcccgatactgatccgcaaggtgtac tcctggtggacgctcgcccATGGTGCTGATCCAGTTCTTCGCCATCCTCGGCAACCTGGCGACGA <u>ACGCGGACGACGTGAACGAGCTGACCGCCAACACGATCACGACCCTGTTCTTCACGCA</u> 20 CTCGGTCACCAAGTTCATCTACTTTGCGGTCAACTCGGAGAACTTCTACCGGACGCTC GCCATCTGGAACCAGACCAACACGCACCCGCTGTTTGCCGAATCGGACGCCCGGTACC ATTCGATTGCGCTCGCCAAGATGCGGAAGCTGCTGGTGCTGGTGATGGCCACCACCGT CCTGTCGGTTGTCGgtatgtgtgtgtgtgtgtggccgtttgggaaagtgtctttgcggcagaaccccaatctactgttacgc ttgactgggtttttgtttttttctcggtggagggatgaaaatatctgaaagaataattgagtcaacccacagggggatgcaag 25 a categorage agage tt t gggtt t gatt tateace geacace gaat at etteace ggtteat aa getteace geggt gaaa agggaleer and the second scctactag CCTGGGTTACGATAACATTTTTCGGCGAGAGCGTCAAGACTGTGCTCGATAAG GCAACCAACGAGACGTACACGGTGGATATACCCCGGCTGCCCATCAAGTCCTGGTATC CGTGGAATGCAATGAGCGGACCGGCGTACATTTTCTCTTTCATCTACCAGGTACGTTG 30 GCGGAATetcctgcgcgtcacagttggcagtcagtgagcggcaacacggcgaaaaaaatgggactaaaaccggtcttcacaga gccaacacattcctacagcaattgcataccttcgggcggtcgggactgggcaatgcagctacaacatcctcgcctaaagttatgcaat tcgagcgacaaatgttgccgtgttagggctttttgtgataatagtcgtttttttgtcctctcgcttatcaaactctatcaacggaggaaatcaacaaattctatgttctcaatggcaaagattactgcccgcaccaatcgcccaacgaaacggcaaaagaaaagcgacgattatga 35 agatgtccaaaccattgcccgcccgacgctttatctgatgatttgcgggatggcttttacttgtctgctactttcaggcacaaaaggaa atgaaaccagcgcaggctcgtttgccggcttgcggaggttcttcaggcactgaggctgagtacttaaatcgaacgatttttacgattc tggatccagttttatgatgtggcctgcattacagtggcaattataccctgatgttcatttcattgcattttgtaagtttgtgctggtaacgcccgtaacgattaattcttttcaaagagattctttcaaagagattcaaaatgtgtataacaaatgctaacgaatggaccgtacttgg 40 gtggcggcagatgtgtcgctgtccgcttccttcctagcaagctcgtgcgaaataatttattccatcattttaatacagccgtttgtg cattttaattagcaaagcaatataaaaagcagctaaccatccccattaaaacaaagtgcttccgggcccaattgttatggcggtgga aagtaatggttttaccagtggaagtgtcctttcccatcgtgggtacttcgcgatattcttgtcttatacaagtgcatacagaaaaaaa ggacaaatcctccttgctatggtctaaggccagcttcggtaccgcttccgcttcgggatgtcataaagtttgatgggtgtttttaacatt acttccgctcttaaccacctaatggacttttcatgcttgagctaaagttaaaccagccaccagcggtacgcaccgagccacggttgatt 45 teggeggeggcctcatccccagttttgcgccaccaatattgccttcattaatctgtaccctcggagcgttagggcccgcggacgagtcct

tgctacaacacattttatgcttcacagatttacttcctgctgttttcgatggtccagagcaacctcgcggatgtcatgttctgctcctggt GAGCTITICGGCCTCGCTGGACACCTACCGGCCCAACTCTTCGCAACTGTTCCGAGCAA 5 TTTCAGCCGGTTCCAAATCGGAGCTGATCATCAACGAAGgtatgtgaaacgtgtgctcgtggcagacg gactcaaagagagcataacacaatcccctggtagttcatttcaatgaccttaacactcggcaagctaagcgagacagtggggacag tgagaaagagagaacaagaaaaaaaccatcatccgtacgacatcatcgctacgtaccggtatttcaggatgaggaaataaaac 10 gagcaaaaaaaagtcaaataaattgaagtttaaaaatagattttccccgtccatccgtggtggagcgtaaagcccggcggacaactt cgagcacggcgaccgtgcacagtactgtgccacagttgtagggacggataagctccgttccttttttatcctttttttggagatttgt catttaatctatcgcgcctgtacgcctgaaactatgcactgtgctgtgaaaccgtcaagctcgagcacgacgaatggcccaccgtacc acgcccgtggtgcccaaagcgcaacgcgaattgcatgttaacaaacctttgcctaccatccaatccgtgtgaaattgcccgctctctt 15 cttaagtccaatgtaatttaccgtgtttctgtcgttcgtcaccttcttcgtcgatggagattggtgcggttggcacgataaaagcccactgcacgttacggaccgagggaaaggtctttttgtaggcctagcaacggtcctcattcaccgcatgggggtgtagctcagatggtagag 20 cgctcgcttagcatgtgagaggtaccgggatcgatacccggcatctccaacccacacaaaacgttttttaagaagatttttagggaa gatattaacgcgggtacactgtgctcctctaagttggaagagtagatgatgatgacaagggagaaggaacatgtgtacgtgttt tgtccctctctctctctgttcaactcctaaaagaattgtttggagtcctctcagttcctcgtaaagatcctttcgagattcttctttttattatttattccacgagcctctgacataagtagccttccgcttatttccttctccttgcacttgtcagttccgtgtagagcgtcattttgag 25 gAAAAGGATCCGGACGTTAAGGACTTTGATCTGAGCGGCATCTACAGCTCGAAGGCGG ACTGGGGCGCCCAGTTCCGTGCGCCGTCGACGCTGCAAACGTTCGACGAGAATGGCAG GAACGGAAATCCGAACGGGCTTACCCGGAAGCAGGAAATGATGGTGCGCAGCGCCATC AAGTACTGGGTCGAGCGCACAAGCACGTTGTACGgtaggtatggtaatttctaaggtgtggtgtaaag 30 ctgaaaccggttgcaatatcgttttgcgaagaaattatgtgtaaagcgtattacaatctcattcctctgttaatctgtaccaattgtgtc agccccgaccgaaagcaggcctaattcgtaccagaaaaaccacaagctgtttgtaagcatcgatacgcccgaagctttcaatccagc cgcccggagtgaagtttttatttgaacgatatcacccgtatcgattttccactaaacatgcttaaatcgtttcacaaagctcccccaaa 35 gagtaaccgaacaacctcttgccgctgcttcacgatatcgaacagcaccaagataagcatccctttttccctagccgatgtctccgata tetegatteegetteeagegaggeaaaaaaggegaaetggetgaceteaceeggggegaggaaaaagegtagggattaegtegagcagcacgagttgtgatttcttcttcttctggttccataaatcgctgacggtttccattaccgcctgcggagtgcacacacgtgaag ggaaagcgaaaacgtttagattccagcagcaacggcagcaccagaagcagcagcagcagcggcaaattgaatcatcctgacgcgat 40 gagttgtctgggttttcgggtcggtggcttacagcaccaccaccatctgctgcagctaatacagctgtaaatttcgttagacatagactt gagccgtgttgctgctggttgcgatacggatcacgtccgattcgattcagcctgcgtgtttttggtgaagatccttatcggtgacccactttaagatagcgataagcttttgatcgaaatagtaaatcaaacattgtttcttttttcctattccaaactgttgccaacctcattattacg 45 tttttgcagcgggtgtatagtaaattgcatactttaaggcgtgattttcaaatgtagcgttccgtatgcagaaacgccatggattatgcaatttaaacaatgctgcttccttaacattcaaataacggcttattaaggaactttttgtgcaatttgttttaacagcaaatagttagc tcagaacgatcacatttagtatcgcttcaacaaagaactcttttaaacacacaatttgtaatgccattccctcgagaaagtttcttgtcagtectectetgeateaeageaacaaceaaacetgeteatgttteetgetegttteetagetgttttgaaegttattteegatteetgtet

tgcccgcttttcttacaatcaaccacaatggttcagatttcgctcttattttattgacccactgctttcgtgctgaagcccgtggaaacaa tgcgccaagctcagcatccagccatgcatgtaaaatgagccacgcgacagattttagacatcgctttcgctctgcaccggaggtggttaacgagccagaaaatgagcacgccaaatgcaaagaaaatccccttttgagtggtgctcctgccaccactcatctccccaactggtgg 5 cg caga aget caa accaa cg ceg ceg cage age at caa caat tte tat te aa acae ce aa cg cage ce caa acceg gg tg caet gt a comment of the commentctcagtagcgaagatgctcagattgtcccgtgcgctgctttcgatgcccgtttcggagcgggaagccatcgcttgccaacgttggcgat ccgtaccgcggtggggcgagttttcaacgcaaccttctacaagcaacgccacaacgcctgggagcgatatttaacagaaacaagaa 10 catecega act teag cae at georgical type category and the context of the contextgggtggcatttgtgtggcatgctatcgtcagcttttcttgaatctttacctctccattcgcctccattagtacacgcgtatggaaaatgg gtgcaacggatcagaacggattttccgcgacagacttaataaagggaaagcaacgcgttttttgcatgtgtagtgtttatgagcttt 15 aaaaaaatgtcaatctgtatcgattattcacacaaatcagatcccggaaccagtgtagcccaatgtgctcttattgaattaccacga gccctcatcgaacagataaacagaagggcaactcttgtgagcatcgcaatgcccgtctgaagttccgtcgaaaatgggcctaaattc 20 tttgggattggtttttgcagcgaaaaatcaaaacattcgcacaaaaccgtcctccatttcaaatgcctacacttgtcactgtatatctct 25 GCTACACATGCTGACCTCCACCATCAAGCTGACGCTGCTCGCCTACCAGGCAACGAAA ATCGACGGTGTCAACGTGTACGGATTGACCGTAATCGGATATTTGTGCTACGCGTTGG CTCAGGTTTTCCTGTTTTGCATCTTTGGCAATCGGCTCATCGAGGAGGAGgtacgtgcgctcggcgtg ttgccgtgggaaagcattctccctgccccatatcgcttcattctcccagatcacacatttgcatcacaaagccagcacacttttgcttcg 30 ccgctgccatctcggcttctgaatgttttcacttctcccatacttctcccgtgcagAGCTCATCCGTGATGAAGGCGGGC $\overline{CTATTCCTGCCACTGGTACGACGGGTCCGAGGGGGGGCAAAAACCTTCGTCCAGATCGTT}$ TGTCAGCAGTGCCAGAAGGCGATGACTATTTCCGGAGCCAAGTTTTTCACCGTTTCGCTCGATCTGTTTGCTTCGgtaagtgtagcctggtggctggcacagaacaggctggcaaaacagggactttggctctagc 35 CTACTTCATGGTGCTGCTGCAGCTGAAGTAAacagccgtggcccggaaggatgtgttttttttcgctcgttcg acagatetttgcaaaatgattagattttaatagattaacagtgettgattatetgteetgtagcaaceggggetgaagaacgttgatt tggtaaaagtacaaaagggacgttggaaattgaaccaccagaagagtgatatttatgcaaagctcaccaagggaaatctatgtat gtgtgatttgcgctcatcaagcactgtatgtgcctttcaactagtgcagcaataaagagtacaaatgtttcttagcgcaccgtacattg 40 45 cagcaaaaaaaatacatataaaaccttcatcactcaagctgtatcgagccagcgtgggttgtgtttgactgtgctgtgaaagaaga agaaaaaaaaaaaaacacttccacgggaagctagcaattggaaatgcataaattaaccggaagaaattcgcaaaaccccgcaccgac

5

Figure 9 continued

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

□ BLACK BORDERS
□ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
□ FADED TEXT OR DRAWING
□ BLURRED OR ILLEGIBLE TEXT OR DRAWING
□ SKEWED/SLANTED IMAGES
□ COLOR OR BLACK AND WHITE PHOTOGRAPHS
□ GRAY SCALE DOCUMENTS
□ LINES OR MARKS ON ORIGINAL DOCUMENT
□ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

IMAGES ARE BEST AVAILABLE COPY.

☐ OTHER: ____

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.